

3.0 **MODEL CAPABILITIES REQUIRED TO MEET USER NEEDS**

Chapter 2 showed how the needs of those who use consequence assessment systems lead to the requirement that an ATD modeling system within the larger consequence assessment system predict accurately and usefully the concentration of the airborne hazard (or hazards) as a function of space and time. The predictions must be relevant to the actual conditions at the time of an incident. Users typically need to know who is not endangered, as well as who may be in danger. They want to know which locations are likely to be within a hazard area, which locations are safely outside the hazard zone, and the quality of these estimates.

Users know that ATD modeling systems do not produce perfect predictions. They desire—and find—ways to work with the limitations of the information, but the modeler's measures and expressions of probability and uncertainty are insufficient to help many users, particularly emergency responders and managers, make sound decisions. This user need imposes two complementary demands on the model developer. The first demand is to provide a reasonable measure of the uncertainty in a prediction or its probabilistic distribution. The second is to communicate the implications of this uncertainty measure or probabilistic distribution in ways the user can apply.

This chapter interprets all of the above user needs into requirements on ATD modeling systems, in terms relevant to assessing the further R&D that should be done. Section 3.1 describes how the temporal and spatial scales for which models have been designed limit their applicability to other scales, either to get input for the model or to apply its results in the real world. Section 3.2 returns to the major functional components of a consequence assessment system, as introduced in chapter 1, to examine how the requirements on the ATD modeling system flow down to requirements on each of its components. It describes current capabilities in each component, compares them with what is required, and identifies both challenges and opportunities in meeting the requirements. Section 3.3 examines, from the standpoint of actions available to the research, development, and test/evaluation communities, ways to

Uncertainty in ATD Model Predictions

The total model uncertainty is measured by the variance between the predicted and the observed quantity over a large number of events that have similar properties (an ensemble). In a recent discussion of the mathematical basis for understanding model uncertainty (Rao 2004), the components of the total model uncertainty are divided into:

- (a) Internal factors such as the numerical approximations to the governing equations, modeling errors, and the treatment of dynamical processes;
- (b) External factors such as data used to execute and evaluate the model, model parameterizations, and the initial and boundary conditions; and
- (c) The stochastic component or inherent uncertainty, due to the natural variability of the atmosphere.

The model developer can minimize the first two components of uncertainty by addressing the several factors contributing to each. The third component cannot be eliminated and is only quantifiable in a statistical sense. Furthermore, we can expect inherent uncertainty to vary as a function of averaging time, location, and the ensemble parameters.

For the analysis of R&D needs, the essential relationship between measurements (observations) and identifying, quantifying, and minimizing model uncertainty must be embraced. The inherent uncertainty cannot be estimated without measurements. Progress toward reducing the first two components of uncertainty also depends on having appropriate observations and on continually improving the techniques used to obtain them.

undertake the task of improving the transition of modeling capability into useful tools for users. In effect, it analyzes capabilities, gaps, and opportunities at the output interface from the ATD modeling system to the consequence assessment system.

The exposition in this chapter draws on two prior reviews of ATD modeling capabilities, each of which included recommendations on R&D needed to address deficiencies. The National Research Council (NRC 2003) reviewed current capabilities in dispersion modeling, identified deficiencies and research needs, and recommended actions to provide more accurate information. The 11th Prospectus Development Team of the U.S. Weather Research Program addressed meteorological research necessary to improve air quality forecasting (Dabberdt et al. 2004). The JAG performed its own survey of current capabilities, which are summarized in Appendix B.

3.1 Consequences of Model Scale

Atmospheric processes are classified by the horizontal dimension and time periods of typically observed phenomena. Choosing an appropriate ATD model requires knowledge of the physical processes that should be treated for the intended application. It also requires an appreciation of the uncertainties associated with the tradeoffs made by the developer in constructing a model of the physical processes that are dominant or relevant at a particular scale.

For purposes of ATD modeling, there are three major scales of interest:

1. **Macroscale** applies to processes having spatial dimensions of 2,000 km or greater and influencing temporal variations of 3 days or longer.
2. **Mesoscale** applies to processes having spatial dimensions of 2 km to 2,000 km and influencing temporal variations of 1 hour to 3 days.
3. **Microscale** applies to processes having spatial dimensions of 2 km or less and influencing temporal variations of 1 hour or less.

These three are further subdivided by decades of distances, from larger to smaller, indicated by α (alpha), β (beta), and γ (gamma), as shown in figure 3.

As the scale becomes smaller, the effects of some processes become increasingly more difficult to treat explicitly or deterministically. Depending on the horizontal scale of interest, different atmospheric processes become significant. Turbulence—the gustiness superimposed on the mean wind—can be visualized as consisting of irregular swirls of motion called eddies. Eddies produce effects at the microscale. The small-scale phenomena associated with the microscale are so transient in nature that deterministic description and forecasting of individual eddies is virtually impossible.

The scales of atmospheric motions are interconnected and nearly continuous. Macroscale processes drive mesoscale and microscale processes as energy is transferred from larger to smaller scales. Conversely, small-scale processes can organize to develop larger-scale systems, such as convective storms. Many of the phenomena of interest for ATD occur in

the troposphere—the portion of the atmosphere from ground level up to approximately 13 km. Most applications of ATD models are for incidents occurring in the atmospheric boundary layer (ABL)—the lowest few kilometers of the troposphere where people live. However, there are situations in which transport and diffusion in the upper atmosphere become critically important for ATD modeling.

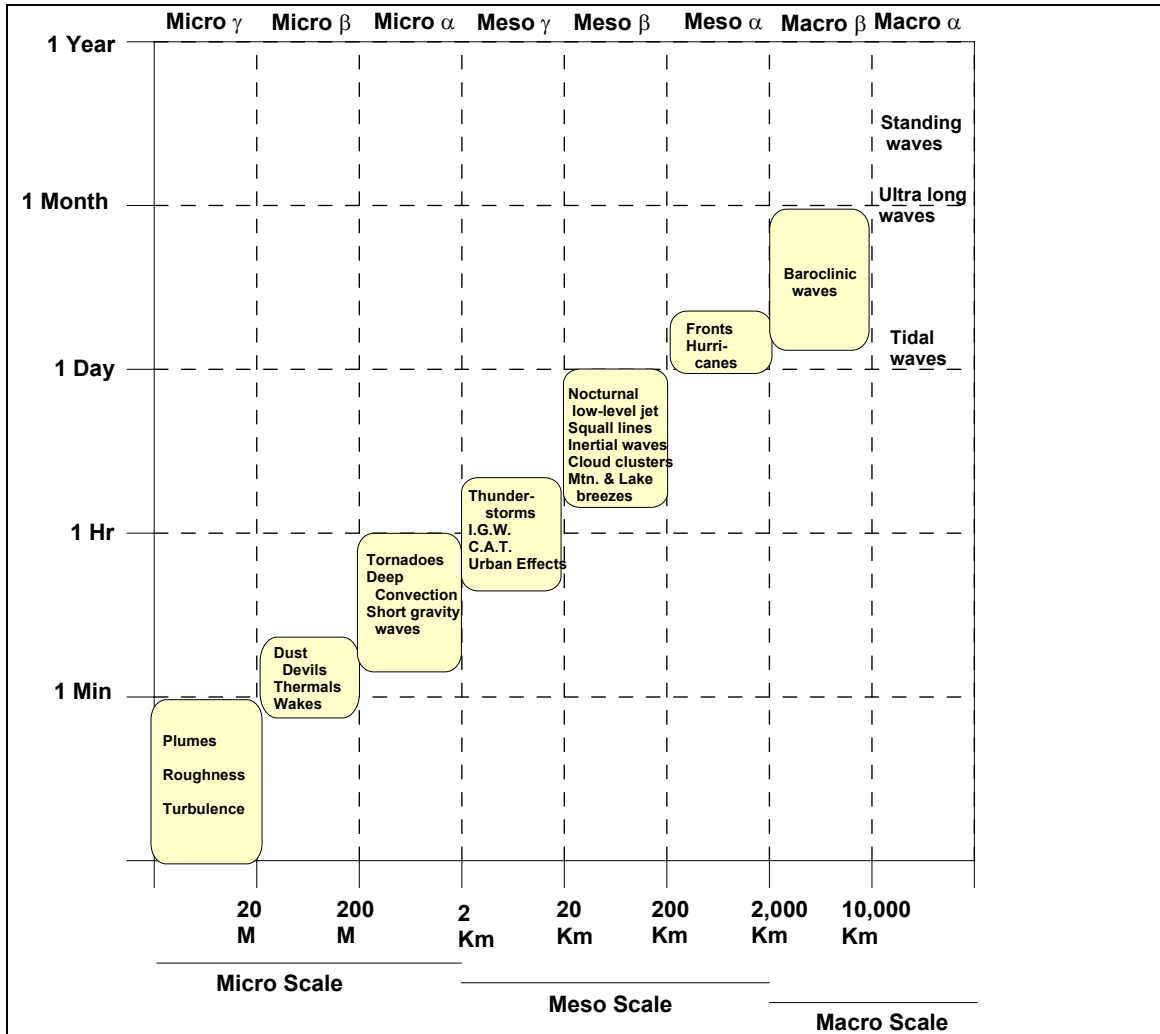


FIGURE 3. Scale definitions and different atmospheric processes with characteristic time and horizontal scales (adapted from Orlanski, 1975). C.A.T is Clear Air Turbulence and I.G.W. is Inertial Gravity Waves.

The horizontal and temporal dimensions of the incident to be modeled define the appropriate scale of the ATD model. The chosen ATD modeling approach should be appropriate for the circumstances, providing a comprehensive and concise description of effects at a particular scale of interest. **Note:** the horizontal grid increment is *not* the scale of the model. Full representation of the phenomena at the desired scale requires five or more grid increments. Appendix C contains a fuller discussion of ATD model construction and selection related to considerations of scale.

3.1.1 Consequences of Scale in Atmospheric Data

Atmospheric measurements may show scaling by their horizontal spacing or by the frequency of observations. Both the spatial and the temporal scale are important to understanding the relevance of observations and their applicability to models. Unlike the continuum of atmospheric motions, measurement scales show little continuity in space.

Table 1 lists common ground-based measurement systems used in the United States and some of their characteristics. The spatial scales they generally represent are also indicated. As the table indicates, the only systems that are truly available nationally are surface weather observations, the rawinsonde upper air system, aircraft data from the Aircraft Communications Addressing and Reporting System (ACARS), and the Doppler weather radar (WSR-88D) system. All of these systems are applicable to measurements of the meso-alpha and meso-beta scale processes. The specialty systems and tracer measurement capabilities are applicable to smaller scales but are available in a relatively few locations and for limited times.

TABLE 1. Spatial Scale and Observation Frequency of Common U.S. Meteorological Observing Systems

Spatial Scale	Observing System	Observation Frequency	Vertical Range	Spatial Separation	Spatial Range	Spatial Resolution
<i>In Situ Measurements</i>						
Meso- α	Rawinsonde	12 hourly	Surface to 30 km	400 km		
Meso- β	Weather observations	Hourly	2–10 m	60 km	Local	Local
Meso- β	Aircraft platform	10 to 1 Hz	Surface to 20 km	Variable	Continental scale over time	Platform dependent
Meso- γ	Tethered balloon	variable 10–30 min	1 km	Irregular	Local	N/A
Multiple	Tracer	1s to 30 min	Local	Irregular	None	Irregular
Micro- γ	Sonic anemometers	10 Hz	Tower height	Irregular	N/A	Tower spacing
<i>Remote Measurements (Excluding Satellite-Based Systems)*</i>						
Meso- β	WSR-88D weather radar	~100 Hz	100 m to > 15 km	200 km	250 km	1 km
Micro- α	Radio frequency sounders	15 min	100 m to >5 km	Irregular	Vertical only	Irregular
Micro- β	Acoustic sounders	10 to 30 s	20 m to 3 km	Irregular	Vertical only	Irregular
Micro- β	Doppler lidar	~500 Hz	~4 km, aerosol-dependent	Irregular	3 to 12 km	3 to 75 m range gate
Micro- β	Radio Acoustic Sounding System	~1 min	100 m to > 5 km	Irregular	Vertical only	Irregular

* Satellite-based observing systems are applicable to many of the scales listed but were not included among remote observing systems in this table.

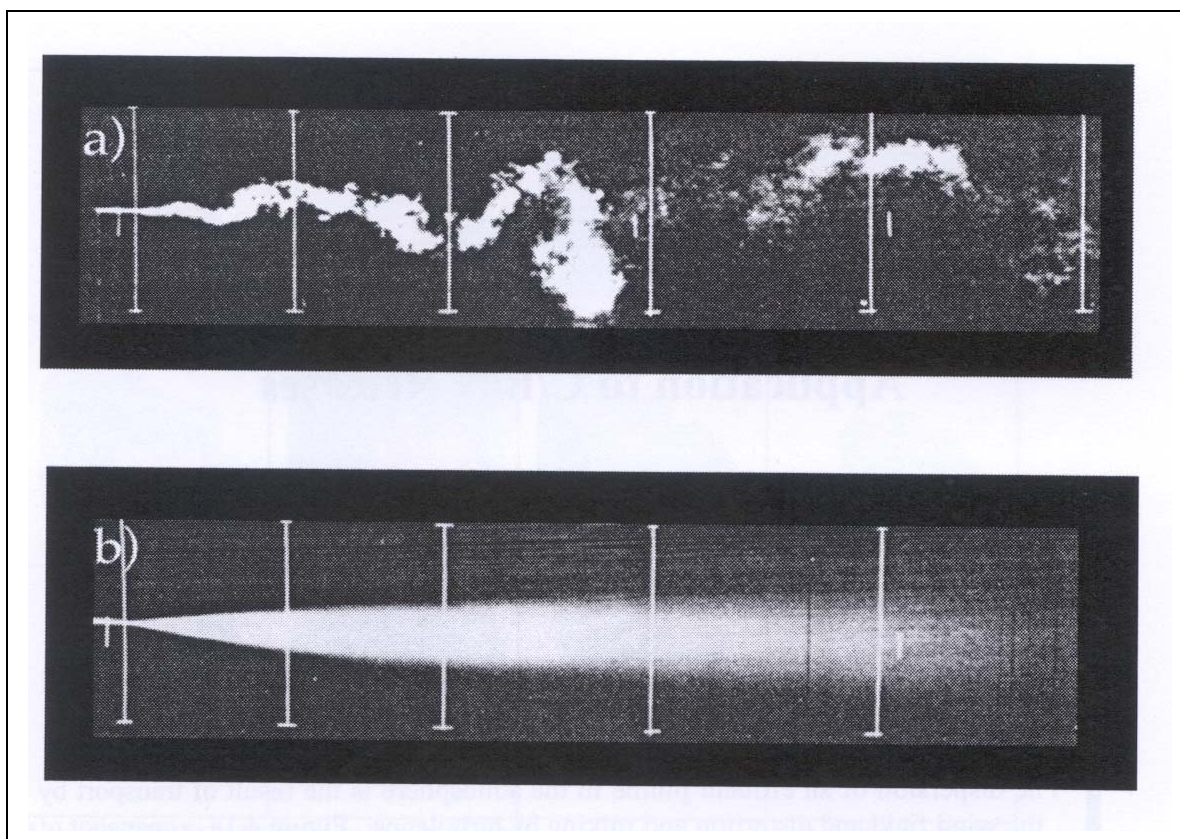


FIGURE 4. Concentration field of a simple flow (top) versus time-averaged distribution (bottom). Two images of the same release. In (a) we see a photograph of an instant during a point source release of smoke within a wind tunnel (view is taken looking down on the plume), where large and small swirls have distorted the plume into serpentine twists and turns. In (b) we see a time-average photographic exposure of the smoke release, where the time-average of the individual chaotic swirls are seen to have the “traditional” Gaussian plume shape used in ATD plume dispersion models. (Photographs are courtesy of U.S. EPA/NOAA Fluid Modeling Facility).

3.1.2 Consequences of Scale in Concentration Data

ATD models attempt to describe hazard zones by their boundaries and temporal extent. The meteorological portion of a model attempts to describe where material would go if the source was known, but the spatial and temporal distributions of the concentration are highly variable. As shown in figure 4, a single realization of a concentration field in a simple flow may bear little resemblance to a time-averaged distribution. At a given location, changes in concentration over time will depend on the sampling frequency of the sensor and its sensitivity. For high sampling rates, a sensitive sensor can detect a few intervals of large values and longer periods of low or no concentration. Depending on application, the time-averaged value may be more relevant or entirely inapplicable.

Short-term peaks in concentration, which are needed to assess acute effects or explosivity, are microscale phenomena. Many other characteristics that affect ATD predictions, such as concentration eddies in the vicinity of walls and urban canyons, are at the microscale. Meteorological models that are used to initialize ATD models are

typically mesoscale models. Issues arise because of the scale differences between the meteorological model's process representations and grid spacing compared with the microscale representations needed by the ATD model.

For long-term health and environmental effects, time-averaged concentration is useful. Wind transport at local scales, however, has a large stochastic component that makes the time-averaged concentration a probability distribution with respect to space and time rather than a point value. To improve the information given to the user, the model researcher-developer needs to represent these stochastic processes realistically in the model and produce a probabilistic prediction that includes measures of the uncertainty in the point estimate (the probability distribution). Then effective ways need to be found to communicate to the user the implications of concentration as a probabilistic function of space and time.

Relevant to the components of an ATD modeling system, these scale issues affect many of the capability requirements and contribute to many of the capability gaps. Discussions of scale will recur repeatedly in section 3.2 and chapter 4, as the specific capabilities, gaps, and the R&D required to address the gaps are presented.

3.2 Requirements and Capabilities by System Component

The functional components of a consequence assessment system, which were introduced in chapter 1, are shown again in figure 5. Each component of the ATD modeling system (within the bold boxes) has its own requirements to become a functional part of the whole. These requirements can be compared with current capabilities in that functional component. Where the requirements are not fully met with existing ATD models (capability gaps), promising directions for further R&D can be identified on a component-by-component basis.

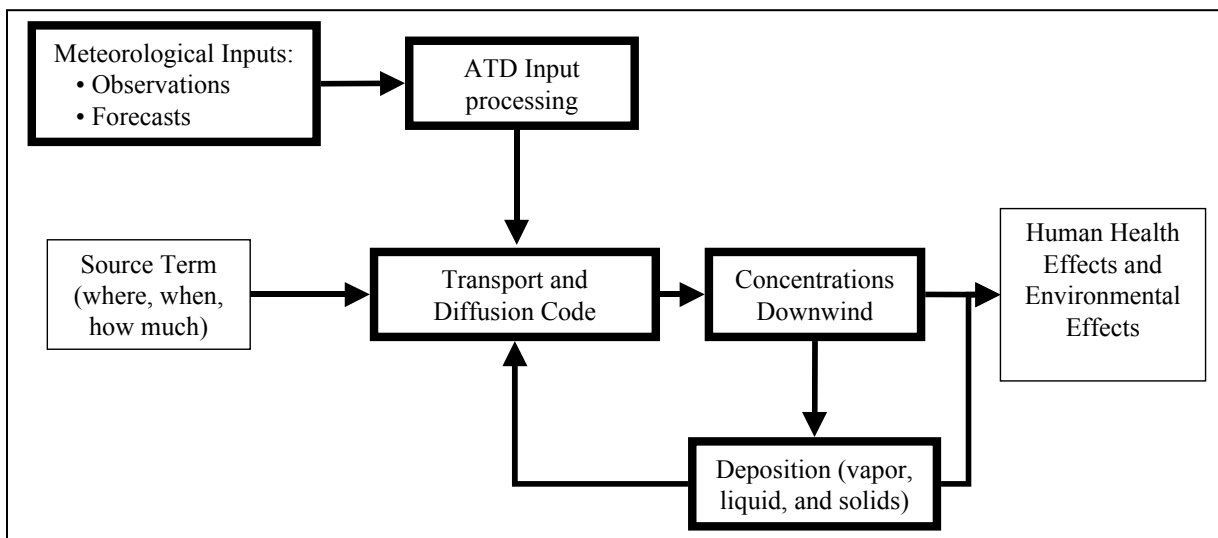


FIGURE 5. The functional components of a complete consequence assessment system. The embedded ATD modeling system is shown by bold lines.

3.2.1 Source Term

The source term component of a consequence assessment system includes information about the identity and physical state of the hazard, the release mechanism, and the mass of hazard released per unit time (emission rate). When ATD modeling is used in emergency situations, the characterization of the source term and the local transport and diffusion conditions are typically the largest sources of uncertainty. For users, the four questions listed in Figure 2 for the release event are source term questions essential to consequence assessment: What was released? When? Where? How Much? The mass of hazard released per unit time, or emission rate, is the key input derived from the source term that the ATD model needs to answer users' questions about where the hazard is going and in what concentration (concentration as a function of space and time).

To characterize near-field (less than 3 km) dispersion, it is critical to know the dilution and buoyancy of the source emissions in the immediate vicinity of the release. Uncertainties in the emission rate and initial dilution volume greatly increase the uncertainty in the near-field impact estimates.

As noted in section 1.2, this report does not address R&D needs for source term characterization. However, ATD modeling techniques can be coupled with concentration measurements made by sensors at some unknown distance from the exact location of the source term to back-calculate to a more precise estimate of the source location and emission rate. This approach, called *sensor fusion*, can be defined for the purposes of this report as the combination and synthesis of information from networked sensors and predictive models to obtain more information about a chemical, biological, or radiological event than would be available from any individual sensor or diagnostic model alone. The networked sensors can include multiple sensor types, including in situ sensors and remote sensors, and other relevant sensors such as meteorological instrumentation. Section 4.4.1 explores sensor fusion techniques and their potential for reducing uncertainty in ATD model predictions downwind from the source location.

3.2.2 Meteorological Inputs

The meteorological inputs to an ATD modeling system may be data from observations, the output from a meteorological model, or a combination of observations and model output. At a minimum, ATD models require wind speed and direction and a simplified turbulence parameter as their meteorological inputs. A more complete specification of the meteorological parameters of interest may include input data on clouds, precipitation, temperature, pressure, humidity, surface heat and momentum fluxes, and a more complex characterization of turbulence. Mass structure and winds can also be measured directly, using a variety of in situ and remote-sensing systems and processing techniques.

In the absence of such detail, ATD models make assumptions to characterize meteorological conditions. Providing as much pertinent meteorological information as possible will improve ATD model predictions by decreasing the number of assumptions that must be made by the model. Before mesoscale meteorological model output was available, ATD modelers used surface and upper air meteorological observations from

sites near the release location. In cases where the nearest available observations did not represent the meteorological conditions at the release site, the modeler would estimate the wind and turbulence conditions.

Mesoscale Models for Meteorological Inputs

While mesoscale meteorological models are executed at much finer grid resolutions than are macroscale (synoptic, global) meteorological models, they typically ingest boundary conditions from a macroscale model. With the sustained growth of computational resources, mesoscale meteorological models that provide acceptable descriptions of mesoscale atmospheric motions and turbulence were developed. These models have now been run operationally for over a decade, and the output from these models is used as input to the ATD model in cases where direct observations of local atmospheric conditions were not available and for the prediction of changes in conditions during transport and diffusion. The use of mesoscale meteorological model output has also allowed ATD model developers to account for additional atmospheric processes with self-consistent input. Although mesoscale meteorological models have proven capable of describing mesoscale atmospheric motions and accounting for atmospheric processes at the mesoscale, they have not yet been optimized for ATD models. Methods to refine these meteorological products before using them as input to the transport and diffusion code component are explored in section 3.2.3.

An advantage to using mesoscale meteorological model output (as opposed to macroscale models) to drive ATD calculations is the potential for improved resolution of localized wind patterns. Worldwide, many population centers are located near coastal regions with highly variable wind patterns. Thermally driven flows associated with land–sea interfaces and complex terrain, which are not resolvable by the coarser grid of macroscale models, can present significant challenges to the accuracy of ATD model predictions. Mesoscale models with horizontal grid lengths of about 12 km or less are capable of capturing some of the time evolution of such flows, potentially improving the accuracy of ATD computations for regions with these wind patterns.

For consequence assessment applications, modeling surface-layer fluxes, winds, and temperatures, even in a mesoscale meteorological model, is a challenge for many regions of interest. Surface fluxes are currently parameterized in numerical weather prediction (NWP) models using Monin-Obukhov similarity theory (Stull 1988). The atmospheric surface layer is defined as the inner region of the ABL, having approximately constant flux with height. It is generally on the order of 10 to 40 meters in depth for neutral to unstable conditions but can be considerably thinner in stable conditions. Because the atmospheric surface layer can be observed continuously using instrumented towers, there is a long history of studies measuring it under a variety of surface and atmospheric conditions. These observational studies have supported the development of detailed theoretical descriptions; however, as originally detailed, these theories are applicable to flat surfaces having uniform roughness, albedo, emissivity, moisture, and thermal conductivity. Real conditions, particularly in populated areas, often deviate significantly from these idealized conditions. So modeling the surface-layer fluxes, winds, and

temperatures in real cases is difficult, even if the larger-scale winds (scales from tens to hundreds of kilometers) could be predicted exactly.

As an example of particular interest to many consequence assessment scenarios, surface irregularities (roughness elements) due to land use (trees, buildings, etc.) are a major challenge for modeling surface-layer properties. Especially in urban areas, large changes in surface conditions (parks, high rises, rivers, industrial zones, residential areas, etc.) can occur within distances of a kilometer or less. This variability affects the local state of the atmospheric surface layer. In major urban centers, tall buildings create “urban canyon” effects. The different types of surface irregularities found in urban areas are difficult to treat in a mesoscale model with a single practical theory for representing the surface layer. In fact, the flaws in current theory for modeling uniform surfaces may be small in comparison with uncertainties due to the effect of spatial surface irregularities found in major urban areas.

Another problem associated with high-resolution mesoscale modeling involves how information is passed from coarser to finer scales when models are nested (a smaller-scale model taking its initialization data and boundary conditions from a larger-scale model). For example, if there is an inconsistency in the nested models’ terrain or urban information databases, errors will propagate to all levels of a simulation. Some models currently in development have two-way feedback, which creates even more sensitivity to the initialization data.

Limitations in Using Model Fields for Meteorological Inputs

Although driving ATD calculations with mesoscale model predictions can, under favorable conditions (i.e., in other than complex environments), improve simulations of transport and diffusion due to localized wind flows, this approach is not without pitfalls. Slight misrepresentations of the temporal evolution (i.e., the timing) of local wind flows can severely degrade the accuracy of the predictions. Predicting the timing of meteorological events, whether synoptic (macroscale) or mesoscale in nature, is one of the greatest challenges in NWP. In these cases, ATD modelers should include phase errors as a contributory source of uncertainty and consider how best to quantify the uncertainty in the prediction stemming from this uncertainty in timing of key meteorologically driven events. Modelers must also have effective ways to communicate the impact of that uncertainty to users; for example, by showing plume development with and without the meteorologically driven event.

Forecast or diagnostic models at horizontal intervals greater than about 300 meters are incapable of explicitly representing ABL circulations, which are dominated by buoyancy and vertical wind shear. In daytime, buoyancy-driven circulations have lateral and vertical scales of the same order as the mixing height, which is typically one to two kilometers. These processes (which are turbulent from a larger view) mix the contents of the ABL. The nocturnal ABL is typically nonbuoyant, and stability resists vertical motion. It is poorly represented in current models because the lateral motion is typically weak, moving material without mixing. Intermittent turbulent events occur almost

without local causes. This extremely complex and poorly understood environment is not modeled with skill.

The atmospheric surface layer occupies roughly the lowest tenth of the daytime mixed layer. Although the atmospheric surface layer is relatively well defined during the day it is less defined during the night. It is a zone of interaction, where heterogeneities in energy, momentum, and moisture dominate ATD processes. Eddy sizes in the atmospheric surface layer are proportional to the eddy's height above the surface. More than half of the energy fluctuations are unresolved. Since the ATD processes cannot be resolved, deterministic models do not apply. Predictions of concentration in this layer are the most important for consequence assessment because this is where human exposure occurs, but they are also the most difficult to make accurately.

The problem of accuracy applies even to relatively simple terrain. Hall and Basara (2004) found that operational mesoscale model predictions of wind speed and directions for the Oklahoma City airport had mean absolute errors in wind speed on the order of 2 ms^{-1} for forecast periods of 6 to 36 hours (figure 6). The mean absolute errors of wind direction were larger than 20 degrees. Other studies of model performance during different seasons and varied terrains found that wind speed errors are typically greater than 2 ms^{-1} and standard deviations in wind direction are greater than 50 degrees (Henmi 2003; Fast 2004). Although operational mesoscale models may have a small bias over many predictions, the predictions for appropriate wind speeds and direction for a given time and place can be expected to differ from concurrent observations.

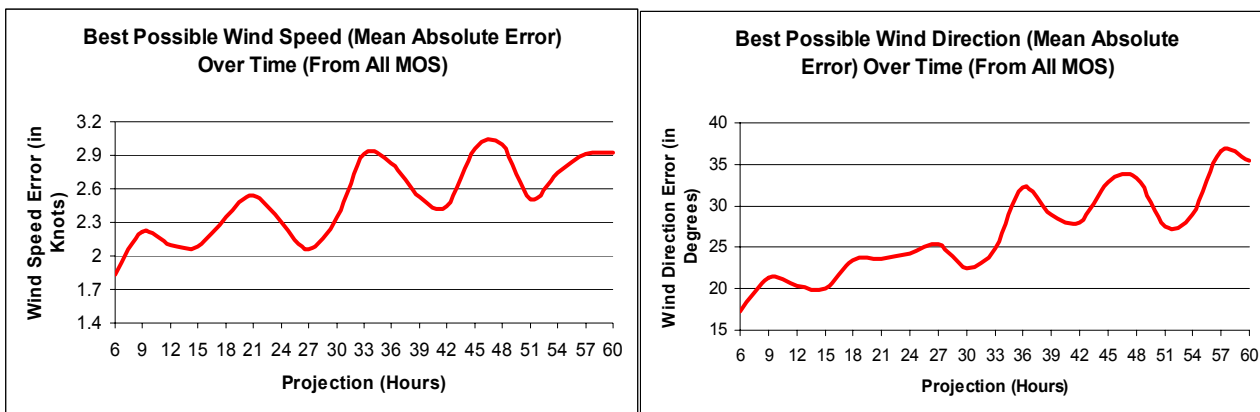


FIGURE 6. Mean absolute error in wind speed and wind direction, measured over a relatively simple terrain. Source: Hall and Basara 2004.

Clouds affect transport and diffusion of airborne materials in several ways. Diminished solar radiation from cloud cover reduces surface heating and convective mixing. Nocturnal cloud cover, even at high altitudes, reduces radiative cooling and influences the development and structure of the stable boundary layers. Insolation also can affect the chemical activity of various agents. Convection in clouds assists the mixing of air above and below the boundary layer: a process that contributes to the dilution of concentration levels at lower levels of the atmosphere. To reduce errors in the prediction of radiative fluxes, cloud information can be assimilated into a model rather than being represented by simple parameterizations. For example, remote-sensing methods can provide cloud

mapping data, including inferred measurement of cloud height, for assimilation into mesoscale models.

Ensemble Forecasting of Meteorological Inputs

To provide some sense of the probabilistic variability of ATD outcomes, it is becoming more common to link ATD models to statistical information constructed from ensembles of mesoscale meteorological models. Means, variances, and correlations of meteorological parameters to be used in the ATD evaluations can be obtained by considering results from the multiple model realizations in an ensemble rather than relying on a single set of point-estimate input assumptions. Ensemble statistics can be obtained by including realizations in the ensemble either from differing models (a multi-model ensemble) or from multiple realizations of a particular model (a single-model ensemble). Multiple distinct realizations from one model can be obtained in various ways such as perturbing the initial conditions, varying the parameterization schemes, using combinations of these first two methods, or varying the grid resolution. Regardless of the ensemble building method, the objective is to characterize quantitatively the range of possible outcomes.

Significant research is still needed in this area. In particular, work is needed to determine the optimal number and types of ensemble members to produce a statistically significantly improved result. Advanced techniques for creating individual members of the ensemble are also of interest. Much development work is needed to link ensemble results from mesoscale meteorological prediction systems to corresponding ensemble systems of dispersion models and to evaluate the resulting probabilistic predictions. Most important, to make ensemble techniques useful to the user, research is needed on how to merge the probabilistic information in ensemble mesoscale meteorological solutions with ATD modeling systems to yield user-tailored probabilistic decision aids.

The WRF Modeling System

With support from multiple Federal agencies, the new Weather Research and Forecasting (WRF) mesoscale modeling system has been developed through an interagency collaboration of the atmospheric science research and operational communities. The National Oceanic and Atmospheric Administration's National Weather Service (NOAA/NWS) is currently preparing WRF applications for operational implementation. The initial WRF system in the High Resolution Window domains will be run as an ensemble of six to eight model versions developed with two dynamical cores, multiple choices of physical parameterizations, and different anomalies in initial and boundary conditions. The number of ensemble members is expected to increase over time with planned increases of computational capacity. By the end of 2005, NOAA/NWS plans to implement WRF at 10–12 km resolution over all of North America. This North American WRF is expected to be replaced by an ensemble system as soon as computational resources allow. The WRF system is designed for applications with grids as fine as 1 km or smaller. Current computer capabilities allow WRF ensembles to be run at high

resolution on regional or subregional domains that are smaller than the *national* weather forecasting requirements.

3.2.3 ATD Input Processing

Processing the input for an ATD model, for purposes of this report, includes refining meteorological inputs, whether from observational or model sources, to prepare them for use in the transport and diffusion code. Input processing techniques attempt to solve (or at least mitigate) several kinds of problems.

One set of problems addressed in input processing comes under the heading of *data representation*. Does a data value, whether an observation from an instrument or a model output value, truly represent the conditions that the model assumes it represents? For ATD modelers, data representation questions such as *Are these data representative?* typically mean *how well do the data meet the assumptions that this model makes regarding the data?*

The second type of input processing problem is *data quality*—how well does a measurement system capture the conditions it is intended to measure? The raw data in the direct signal output from a sensor can be in many forms. Most common signals are in the form of an electrical impulse, voltage, current, or resistance or a change in one of these properties. Quantifying the physical principle of the signals to a concentration, wind velocity, or pressure is the first step in ensuring data quality and is part of the sensor design. Data quality also depends on the sensitivity of the physical property it is intended to measure, changes in that property, and confounding environmental conditions. Calibration of the mean and variance of the measurement instrument to known references sets the precision of the measurement.

When each sensor is well calibrated and working properly, the measurement system as a whole may or may not be providing a realistic “observation” of the patch of reality it is intended to observe. At the level of accepting a set of data values from a measurement system, data quality acceptance/quality control (QA/QC) processes may identify outliers as potential instrument errors, interpolate for lost or missing data, or compensate for timing errors or irregularities. Furthermore, the exposure of the instrument and the heterogeneity of the instrument location must be factored into the assessment of acceptable data.

An ATD model generally assumes a correlation or coherence among the input data. In some instances, data incoherence arises from a data quality or data representation problem; in other instances, it results from incompatibility of different input sources (observations and forecast models). When those data do not support the coherence assumption, the ATD model must provide rules for acceptance or rejection.

Development of guidelines for observation networks is one of the R&D needs that emerges from issues of data quality and data representation. In some ATD modeling systems, evaluating input data for either data representation or data quality is incorporated in the input processing operations. In other modeling systems, these

characteristics must be examined independently. Data quality and data representation concerns are further magnified when the data are used to calculate derived quantities such as fluxes, scaling parameters, mixing height, wind shear, or thermal stability. The user of the input data needs to know the temporal or spatial averaging that has been used to produce the derived quantity.

Data assimilation is another concern in processing data for input to an ATD model. It overlaps with data representation and data quality but can also derive from other complications in the data–model interface. Weather forecast models are re-initialized at regular intervals using previous forecast fields and recent observations. In some instances, large differences between the forecast and the observation may occur. The initialization procedure is designed to weigh the forecast field and the observation within the context of expected variance in the values and the governing equations of motion. The new initialization may not include, or assimilate, the observation because doing so would violate other model constraints. This rejection of the observations by the model’s rules for assimilating data may result from data representation problems (data that are not representative of scales that the model can represent), data quality problems, model errors (representations or parameterizations that deviate from the real processes being modeled), or a combination of these factors. The data quality and representation problems can be either in the data the model is now trying to assimilate or in data previously used for initialization or assimilation.

As this limited discussion illustrates, ATD input processing quickly becomes complex, consuming both time and resources (computational and human capital). While automated input processing is appealing, one approach does not fit all models or even all circumstances for the same model. As new instrumentation is developed, a major concern should be for internal consistency checks and usability (suitability) of the measurements in data input processing.

3.2.4 The Transport and Diffusion Code

The transport and diffusion code describes in algorithms the combined effects of time-averaged transport (which has traditionally been viewed as a deterministic process) and of atmospheric diffusion (which has traditionally been represented as a stochastic process). The entire set of computations is sometimes called the “ATD model,” but that term is also sometimes used to mean the way in which transport and diffusion processes are represented when the set of instructions (the code) is run with an appropriate set of initialization data.

Federal agencies, the academic community, and others employ a large number of ATD modeling systems for a variety of purposes, including regulation, research and development, and emergency operations.¹ However, the JAG/SEATD report identified only a few basic types of operational ATD models, or transport and diffusion code approaches, in the modeling systems assessed by the JAG/SEATED (OFCM 2002, p. 1-

¹ The JAG/SEATD reported that the FEMA Insurance and Mitigation Administration identified more than 140 ATD modeling systems in an internal report (OFCM 2002, p. 1-2).

4). These basic model types—box, plume, segmented plume-puff, Lagrangian particle, Eulerian grid, and computational fluid dynamics (CFD)—are defined and discussed in Appendix C. There are also only a few types of diffusion characterizations in use; the most common are empirical, statistical, similarity, eddy diffusivity, and second-order closure. The profusion of ATD modeling systems arises from the variations and combinations of these approaches combined with specializations made to handle unique problems, such as plume impaction on elevated terrain, concentration within the wakes of buildings, or heavy-gas effects.

Because of the stochastic component, all ATD modeling with a transport and diffusion code must be considered a forecast of possible outcomes. In addition, the sets of deterministic and probabilistic equations implemented in any given model provide only an approximation to the complex atmospheric conditions the model is meant to represent. Consequently, ATD modeling is always a compromise between getting a useful solution in an appropriate amount of time and realistically portraying the transport and diffusion of a released material within the atmosphere. These uncertainties introduced by the inherent probabilistic nature of the processes and by the compromises to make the model useful are in addition to the uncertainties in the input data.

Several techniques are used to apply ATD models to complex environments such as cities or coastal areas. The top-down approach uses multiple nests of finer-scale models within coarser grids to approximate the mean transport and turbulent flow at short temporal and spatial scales. This approach is useful when appropriate observation systems at the smaller scale are lacking. The bottom-up approach uses physical models—based on wind tunnels or flow channel experiments for example—or high-resolution computational models (discussed in section 4.2) to capture the larger-scale effects of the complex region being modeled and the fine-scale features of flows within the region. Physical models are used principally to build a knowledge base about a specific location and to provide appropriate data for improving the understanding of processes that are not measurable in the natural environment. Their advantage is that experimental conditions can be controlled; their disadvantage is that each experiment is only one possible realization of the stochastic variability. A middle ground, computational fluid dynamics (CFD) models, uses CFD codes adapted from the aerospace industry for examining turbulent atmospheric flows around single or multiple obstacles. In a sense, CFD models are numerical surrogates for wind tunnels or flow channels.

An emerging option for model refinement is to use remotely sensed data of actual conditions. Remote sensing can inform and update a model of the physical landscape on a recurring basis, allowing natural and manmade changes to be incorporated.

Finally, one of the principal user needs identified in Chapter 2 is seldom met at present. Most of the current operational ATD modeling systems for consequence assessment in civil emergency response applications are unable to provide information on the variability of hazard concentration on the shorter time scales needed to assess such consequences as acute effects of exposure or explosivity. A few modeling systems attempt to estimate the probability that such events could occur but are not specific as to when or where. Even CFD estimations cannot predict the exact stochastic pattern of

dispersion. As a result, a forecast from even a very sophisticated ATD model has a large single-event uncertainty. At present, even ensemble-based ATD modeling systems predict only the ensemble-average dispersion pattern (the average over the multiple realizations in the ensemble) and the range of predicted ensemble variables, not the complete event-to-event variability. Because this variability can represent substantial uncertainties with respect to human health and safety risks, the ATD R&D community must do better at quantifying the uncertainty and communicating its implications effectively to those making emergency response decisions (or other decisions based on assessing consequences sensitive to this uncertainty).

3.2.5 Deposition (and Other Removal Mechanisms)

Substances released into the atmosphere will stay there, continually dispersed and diluted by mixing processes, until they are removed by reactions with other components of the atmosphere or are deposited on the Earth's surface. Consideration of in-air reactions is essential in modeling the ATD of gaseous and biological agents. Nerve agents, for example, interact with atmospheric oxidants and with other constituents of the background air, gradually reducing the total amount of the hazard remaining in the air. Biological agents tend to be susceptible to ultraviolet radiation; hence, their active residence time in the air is largely controlled by their exposure to sunlight. Atmospheric reactions during transport can also be important for reactive liquids and solids.

Deposition Mechanisms

Precipitation is one of the most efficient mechanisms for removing pollutants and other substances from the air. The two precipitation-related processes of importance are *rainout* and *washout*.

Clouds serve as dynamic systems for processing air that passes through them, concentrating most pollutants in cloud droplets, which then coalesce and eventually fall to the surface (ground or water) as precipitation. This process of in-cloud scavenging is commonly referred to as "rainout." For rainout to be an efficient removal mechanism, the hazardous material or pollutant must become directly entrained into a cloud. The scavenging efficiency depends on the chemical and physical properties of the pollutant in question, as well as on the dynamic characteristics of the cloud. Not all of the materials entrained in cloud circulations will be removed and deposited in precipitation on their first pass through a cloud. Sulfate particles, for example, are likely to pass through several clouds before being scavenged. For the gaseous and biological warfare agents of current concern, rainout appears likely to be important but has not been extensively studied.

Hazardous materials that are dispersing in the air near the surface will be scavenged by raindrops or other hydrometeors such as snowflakes, in addition to any rainout scavenging by clouds from which the precipitation derives. This process, called "washout," is relatively inefficient for liquid or solid hazards unless the particles being scavenged are close to the size of the droplets scavenging them. Gaseous hazards, if they

are soluble in water, can also be removed by falling hydrometeors, sometimes quite efficiently.

Dry deposition to the surface continues at all times, regardless of whether precipitation is occurring. Dry deposition is a far less efficient process than wet deposition but often removes similar amounts of material solely because the process is continuous, albeit slower.

The amount of deposition as a function of space and time is complex and difficult to predict in detail. For example, the factors of timing, amount, and location of precipitation are very important for wet deposition of dispersing materials. Prediction of clouds and precipitation mechanisms are a major focus of high-resolution mesoscale models. Although the best current models still have problems predicting the location and intensity of precipitation at scales of interest to potential users of the predictions, they do better when the driving forces are strong. Quantitative data on cloud-mixing processes and deposition are needed but are difficult to obtain.

Descriptions of deposition processes, particularly quantitative descriptions, need to be refined and tested. In future field studies and experiments on ATD, a component to measure deposition rates should be included wherever possible. This necessary work can build on the long history of relevant studies.

Resuspension

The arrow in Figure 5 from the Deposition box back to Transport and Diffusion Code represents the resuspension of hazardous material particles. For ATD modeling purposes, resuspended particles can be treated in the ATD model as a new release or emission to the atmosphere. In some instances, deposited materials will remain at the surface, with potential for subsequent resuspension into the air. Resuspension can be a major consideration for consequence assessment; radioactive particles in surface dust are a good example. In practice, resuspension of deposited materials will occur only when mechanical or volatility forces on the deposited material are sufficiently energetic. Such forces may be associated with vehicular traffic, foot traffic, or simply the wind.

Health and Environmental Consequences of Deposition

Atmospheric deposition provides the linkage between air concentrations of hazardous materials and surface environmental consequences. Although deposition constitutes a major *sink* for removing airborne hazardous materials, it is also a major *source* for studying and assessing environmental effects of hazards. Hazardous substances deposited from the air to the underlying surface are likely to enter into the biosphere. For example, if a nuclear or radiological material were deposited on the ground and inserted in an environmental pathway that led to human food sources, there would be human health consequences from this route of exposure.

3.2.6 Concentrations Downwind

Prediction of the concentration of a released hazardous substance as a function of space and time is the reason why consequence assessment systems incorporate an ATD modeling system. Once all the *appropriate* information about “concentrations downwind” has been delivered, the ATD modeling job is done; other components or players take that information as input for assessing the human health and safety consequences and environmental effects. A major theme of this report, however, is that deciding what information about concentrations is appropriate is not the sole province of either the ATD modeling community or the user community. Much work remains to be done by both communities to meet the user needs set forth in chapter 2.

3.3 Transitioning New Research and Development Capability to Operations

The term “operations” refers to the application of ATD prediction capability by a user to support that user’s decision-making process. As discussed in chapter 2, consequence assessment tools are designed to support a range of operational planning, response, and recovery efforts. The ATD modeling system is likely to be only one component within a larger system for the overall consequence assessment. Transitioning ATD codes or systems from development to operations requires an understanding of the operational requirements, as well as how the ATD prediction capability will be used and how it will be integrated into the larger concept of operations.

Experience has proven that hazard assessment and decision information must get to the *right people* at the *right time*. The “right time” means that information must flow to the decision maker before it is too late for the mitigating action to be relevant. The “best” hazard analysis, if too late, is useless for response decisions, although it may still be relevant to forensic analysis during recovery activities. In addition to timeliness, the information must be operationally relevant.

For new ATD prediction capability to be successfully transitioned from R&D to operational use, the following areas must be addressed: usability; training; data connectivity; results communication; operational testing and evaluation, including production readiness; and documentation. Each area is discussed separately below, but there are major interrelationships among them that are critical to successful R&D. The successful program manager applies sound risk-management processes to invest in and coordinate activities in these areas. Keeping in mind that risks range from low probability of occurrence to high probability of occurrence and from small consequence to huge consequence, it is clear that the risk management plan must describe the risks to the program and prioritize them by degree of importance to the success of the program. The risk management plan should address all of the applicable risks, including acceptability; schedule; and technical, cost, and program risks.

The task of development is not complete until the new capability has been proven useful in operations. The work of the researcher must be guided by what users need and by what

current capabilities cannot give them. To transition a new capability into operations in the time desired to meet national goals of preparedness, upfront and continuing interactions between users and researchers-developers must replace the leisurely, phased approach to research, followed by development, followed (perhaps) by operational deployment. No longer can the researcher or the developer walk away from the issues of transition as being someone else's problem.

3.3.1 Usability

Usability refers to the relationship between tools and their users. An effective tool allows the intended user to accomplish a given task in the best way possible. For ATD model codes that are either new or modified as a result of new research, the intended users should be clearly stated. As the level of user expertise with predictive modeling codes moves away from trained meteorologists and dispersion modelers, the need increases for more complex intelligence to be built into the modeling system to guide the user. For example, both novice users and advanced but infrequent users will probably need simple graphical user interfaces with standard defaults. More-expert users will want to use shortcuts and have more control over input parameters. Emergency response use will generally require a model that adapts to quickly changing conditions, provides clear guidance on input, and allows for unambiguous output. Regardless of user expertise, on-line help and error and range checking embedded in the modeling system software should be part of any operational system that will be used under stressful conditions.

In using dispersion models for planning or post-event analysis, the user friendliness of the modeling system is generally less critical. In nonemergencies, the more flexible time scale for providing an answer typically allows the user to analyze input and output more closely, get additional expertise or data, and explore a broader range of scenarios.

Without a clear understanding of the intended model use, the model user, and how information must flow to get relevant information to the right people at the right time, research-derived model enhancements will fail the usability requirement for transitioning to operations. Proper usability testing and implementation is critical for ATD models designed to define hazard areas where lives may be in danger. Usability testing should address a number of factors including fitness (how well the functionality fits the user need), ability to perform the intended task correctly, and how well the application fits the user expectations. Achieving this level of usability requires iterative interaction between users and developers, beginning well before a modeling product is ready for operational testing.

Prototyping can be an effective means to manage this and other technical risks. The user's inputs should be incorporated in the design of ATD modeling system improvements, and user feedback should be incorporated in subsequent prototype development cycles. In considering tradeoffs between capability and cost, a sensitivity analysis of new approaches or parameters should be part of the prototyping effort. The preferred software engineering methodology incorporates risk management techniques

and engages users and other stakeholders throughout the software system life cycle. An example of such a methodology is the spiral model.²

3.3.2 Training

The model end user is rarely the model developer. Training of both the person who runs the model and the person responsible for making a decision using the model's output is critical for appropriate model use. Unfortunately, the decision makers often do not have the time and resources to be trained to use every tool intended to help them make a decision. Therefore, it is even more critical that the person running the model understand how to convey the implications for the decision maker of a forecast from the model (or from a set of models, depending on the user). As model forecasts become more sophisticated—for example, by incorporating reasonable and useful measures of uncertainty—the forecast itself must be presented in ways that are immediately meaningful to the decision maker. Model developers can no longer rely on the expertise of the person running the model to interpret this complex, sophisticated information and convey it concisely yet correctly to the decision maker. This means the developer (and behind the developer, the researcher) need to be “trained” on the user's decision-making environment just as much as those who run the model or make decisions using model output need training on the tool. In effect, the model must talk the decision maker's language. Therefore, those who create the modeling capability must also understand and “speak” that language.

Analogous to forecasting the weather with a meteorological model, any given ATD model has strengths and weaknesses, depending on the scenario and the environmental conditions known at that time. The forecaster needs to understand the model and the scenario details well enough to know how to adjust the forecast product. Unlike meteorological models that are run daily, thereby generating forecasts that can be evaluated every day, those who run ATD models are often intermittent users. They seldom have adequate data to evaluate the model or enough experience to make reasonable adjustments to the model output. Infrequent model use creates a unique set of problems, some of which can be addressed by usability in the model development. Others can be addressed through appropriate training. Training must address the entire range of users for whom the modeling system is intended to be an appropriate tool.

Although there are a variety of users, most operational objectives share a common requirement—generating consequence assessment information. At present there is no overall certification process for training personnel in ATD modeling. The most-effective training will cover more ground than just using a given model. It may, for example, include learning about the operational environment, exploring the basics of how air moves particles, understanding forward deployable technical solutions and expert reach-back services, and learning strategies for managing the risks of CBRN hazards. Workshops, formal courses, computer-based or on-line training, and tutorials are all mechanisms for providing training that should be considered when new ATD prediction

² The spiral model was initially described by Barry Boehm in 1988. It is a “risk-driven determination of process and product; growing a system via risk-driven experimentation and elaboration; and lowering development cost by early elimination of nonviable alternatives and rework avoidance” (Boehm 2000).

capability is being transitioned into operations. In chapter 5, the advantages of test beds for ongoing and interactive training of both the users and the researcher-developers will be emphasized as the “hands on,” experiential learning to complement these conventional approaches to user training.

3.3.3 Connectivity to Data Sources

Figure 5 identifies the major components of an ATD modeling system as meteorological inputs and input processing, transport and diffusion code, and deposition (fate of the dispersed material), with concentrations downwind as the output. The larger consequence assessment system includes source term characterization and effects on human health and safety and the environment from the dispersed material. All of these components may contribute data to the ATD model. The ability to connect to different data sources for inputs requires an information infrastructure to answer such questions as:

- Are the data available?
- Through what mechanisms are they available?
- What are the temporal and spatial scales for data retrieval?
- Are there standard formats?

Data connectivity also assumes an understanding of how the model will use the data input. An ideal operational system will have a seamless mechanism for both inquiry about access to potential data sources and utilization of the data received by the model.

3.3.4 Results Communication

Requirements for an operational ATD modeling system to communicate a forecast of hazard zones will depend on whether the forecast is for planning, response, or recovery (including post-event assessment). The situations with higher stress for users and less flexibility in timeliness of decisions require more emphasis on standardized, easy-to-interpret output. In emergency response, for example, standardized products for similar categories of threat (radiological, biological, chemical) will aid in the time-critical use of predictions. Planning and post-event assessment provide more opportunity for discussion and alternatives for presenting model output. Whether output is deterministic (a single best guess), probabilistic (probability distribution), or ensemble (combinations of different model outputs), communicating what the particular output conveys and its associated confidence or uncertainty should be considered integral and essential features of an operational system. In addition, an operational capability should provide interfaces for both the most widely available and the latest technologies for communicating output.

3.3.5 Evaluations of Modeling System Performance

This document uses “modeling system performance evaluation” to refer to a collection of engineering and scientific processes that enable modeling system developers to establish the degree of correctness of the software, how well the physical models and databases

represent reality, and the fitness for use of an ATD modeling system. There are established guidelines and consensus approaches for evaluating ATD modeling system performance that must be incorporated in the overall processes of system development, evaluation, and transition to practice, especially when the ATD modeling system is integrated into a consequence system.

The manner in which a modeling system performance evaluation is conducted should depend on a number of factors, including the intended application, whether the modeling system will interface with a mission-critical system, and the amount and type of evaluation processes that were previously applied to the parts of the ATD modeling system. The processes in a modeling system performance evaluation include:

- **Science peer reviews.** During science peer reviews, the model's key constructs must be shown to be reasonable and defensible for the defined uses. A key part of the scientific peer review will include the comparison of modeled and observed evaluation objectives over a range of model inputs (e.g., maximum concentrations as a function of estimated plume rise, stability, or distance downwind).
- **Diagnostic and performance evaluations.** Diagnostic and performance evaluations are two types of statistical evaluations that are typically performed to assess different qualities of how well a model is performing. Both are needed to establish credibility within the client and scientific community. Diagnostic evaluations examine model capability to simulate individual processes that affect the results (e.g., droplet fall velocity using small-scale data sets, such as those from special field experiments, wind tunnels, or other laboratory equipment). Performance evaluations, particularly those conducted in circumstances of the intended application; enable one to decide how well the model simulates the average temporal and spatial patterns seen in the observations. Work is underway to develop a new generation of evaluation metrics that takes into account the statistical differences (in error distributions) between model predictions and observations.
- **Supportive analyses** (e.g., software verification, sensitivity, and uncertainty analyses). Software verification is the process of determining that a model implementation accurately represents the developer's conceptual description and specifications. These supportive analyses should be applied to ensure that the following four key tasks are completed:
 1. Modeling assumptions, limitations, and errors are adequately documented.
 2. The software development effort is well managed and controlled.
 3. Results produced by the modeling system are stable and predictable.
 4. The results of diagnostic and performance evaluations are well understood.

In summary, numerical comparison of model predictions with observed field data provides only a partial means for assessing model performance. Due to the limited supply of evaluation data sets, there are severe practical limits in assessing model performance. In this context, conclusions reached during the scientific peer reviews and the supportive analyses have increased significance in deciding whether a model can be applied in the

circumstances defined by the model evaluation objectives. Therefore, setting up an evaluation program might include publishing peer-reviewed papers, hosting technical review boards, and having independent third-party reviewers.

3.3.6 Software Testing and Evaluation Including Production Readiness

When the ATD modeling system is integrated as a component system of the larger consequence assessment system, software testing must be conducted at all phases of the modeling system's life cycle, starting with unit-level testing and continuing through systems-integration testing. These activities provide confidence that the modeling system's *performance requirements* have been met and determine the degree to which the modeling system represents the real world *in the context of the intended use of the model*. The JAG/SEATD report reviewed the procedures currently in use by Federal agencies for testing and evaluating ATD modeling systems (OFCM 2002), and those procedures need not be reviewed again here.

Established test and evaluation procedures, including the model performance evaluation processes discussed in section 3.3.5 and their documentation, are essential parts of the process of transitioning from an R&D result to an operational tool. Implementation of new research results into new and existing ATD modeling systems should ensure that the following conditions are met:

1. New products of research should make a measurable improvement in and increase the value of the model results to the end user.
2. Software verification and validation procedures should be employed to ensure that new algorithms and techniques perform as intended. If the modeler and the researcher are not the same, then the model developer needs a mechanism to confirm that the new enhancement is being correctly implemented.
3. Usability testing has been completed, and the modeling system meets the needs of all its intended users. Operational test and evaluation should focus on the operational effectiveness of the system and its suitability for operational use.
4. Production readiness has been achieved by demonstrating reliable, sustained production. Production readiness also includes providing results within the required time constraints and providing backup against single points of failure in production, communication, and connectivity.
5. Comparisons with field data have produced no surprising discrepancies. To the extent model results are available, they should be compared with field data or historical data sets. While there will often be differences, the evaluation should be able to explain why the differences are acceptable.
6. Model-to-model comparisons are consistent for different modeling systems used in operations. If implemented correctly, new research will lead to model advances from the private, military, and public application sectors of the R&D community. Testing of multiple model implementations can provide valuable insight on potential problems.

3.3.7 Documentation

Public ATD models should have a range of documentation available:

- User documentation with point-and-click details for the intended user;
- Technical documentation so that other researchers or model developers can independently evaluate and test specific algorithms; and
- Quality assurance and testing documentation.

The code or modeling system should not be considered operational without these documentation components.

